

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

ГУРМАЧ ВАСИЛЬ ВАСИЛЬОВИЧ

УДК 577.322.23+616.379 – 008.64: 616.85

ПОШУК *in Silico* ЛІГАНДІВ ДЛЯ ДОМЕНІВ SH2

03.00.02-біофізика

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня

кандидата біологічних наук

Київ–2016

Дисертацією є рукопис

Робота виконана на кафедрі біофізики Навчально – наукового центру (ННЦ) “Інститут біології” Київського національного університету імені Тараса Шевченка МОН України

Наковий керівник: доктор фізико-математичних наук, професор
Прилуцький Юрій Іванович,
ННЦ “Інститут біології” Київський національний університет імені Тараса Шевченка МОН України, професор кафедри біофізики.

Офіційні опоненти: доктор біологічних наук, професор,
член-кореспондент НАН України
Говорун Дмитро Миколайович,
Інститут молекулярної біології і генетики НАН України, заступник директора з наукової роботи;

доктор біологічних наук,
старший науковий співробітник
Данилович Юрій Володимирович,
Інститут біохімії ім. О.В. Палладіна НАН України,
старший науковий співробітник відділу біохімії м’язів.

Захист відбудеться “26” “квітня” 2016 року о 16⁰⁰ годині на засіданні спеціалізованої вченої ради Д 26.001.38 Київського національного університету імені Тараса Шевченка за адресою: м. Київ, просп. Академіка Глушкова, 2, ННЦ “Інститут біології”, ауд. 434

Поштова адреса: 01601 м. Київ, вул. Володимирська, 64/13, Київський національний університет імені Тараса Шевченка, ННЦ “Інститут біології”, спеціалізована вчена рада Д26.001.38.

З дисертацією можна ознайомитись у Науковій бібліотеці імені М.Максимовича Київського національного університету імені Тараса Шевченка за адресою: м. Київ, вул. Володимирська, 58.

Автореферат розісланий “__” _____ 2016 р.

Вчений секретар
спеціалізованої вченої ради,
доктор біологічних наук

Т.М. Фалалєєва

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Сукупність протеїнових структур формує складну мережу взаємодій, під час яких відбувається передача сигналів як всередині клітини, так і між ними [Henry D.H. та ін., 2013]. Протеїн-протеїнова взаємодія (PPI) – це постійний або тимчасовий фізико-хімічний вплив протеїнових структур одна на одну (утворення/руйнування високомолекулярних комплексів). Ця взаємодія відбувається завдяки різноманіттю протеїнових доменів [Sara C. та ін., 2006], які входять до складу протеїнових структур, характеризуються компактністю [Richardson J.S. та ін., 1981] та широким спектром функцій [Bork P. та ін., 1991]. Так, наприклад, домени SH2 входять до складу протеїнкіназ (Src, Lck), фосфатаз (SHP2, SHIP2), фосфоліпаз (PLC γ 1), факторів транскрипції (STAT), регуляторних (SOCS), адаптерних (Grb2) і структурних протеїнів (SHC). Враховуючи таку розповсюдженість протеїнових доменів, їх розглядають як зручні терапевтичні мішені [Smithgall T.E. та ін., 1995].

Водночас, пошук *in silico* нових сполук зі специфічною біологічною дією є складним завданням, яке потребує використання сучасних методів молекулярного моделювання [Hurmach V.V. та ін., 2012], достатньої кількості структурних і біологічних даних щодо мішені (підтвержені активні та неактивні речовини, відомі комплекси протеїну мішені та ліганду).

SH2 – компактний глобулярний домен, що містить близько 100 амінокислот, які формують дві α -спіралі (α A, α B) та сім β -складок (β G, β B, β C, β D, β D', β E і β F). Його основна функція - взаємодія зі структурами, що містять фосфорильований залишок тирозину (pTyr). Здебільшого домени SH2 містяться в онкопротеїнах (Src oncoprotein) та протеїнах, які входять до сигнальних каскадів клітини. Встановлено зв'язок доменів SH2 з процесами, пов'язаними з онкозахворюваннями [Lappalainen I. та ін., 2008]: Наприклад, порушення в їх функціонуванні викликають гострий лімфобластний [Lacronique V. та ін., 1997], хронічний мієлоцитарний [Peeters P. та ін., 1997], та ювеналій мієломоніцитарний [Tartaglia M. та ін., 2003] лейкози тощо. Важливо, що для цієї родини доменів відомі зв'язані з лігандом або іншими протеїнами структури (наприклад, PDB – 1O49, 2GE9). Саме тому поміж великої кількості доменів PPI, як мішень, обрали домени SH2.

Домени SH2 можна поділити залежно від специфічності розпізнавання pTyr-залишку із C-кінця: таке розпізнавання відбувається у позиціях +1, +2 та +3 [Smithgall T.E. та ін., 1995]. Отже, можна припустити, що кожен окремий домен SH2 зв'язується лише з конкретним pTyr-фрагментом. Наприклад, Src SH2 переважно розпізнає Glu-Glu-Ile (взаємодіючий фрагмент pYEEI), а Grb2 SH2 зв'язується з іншим фрагментом – pYVNV. Однак повне розуміння цього ефекту потребує детального вивчення термодинамічних особливостей взаємодії фосфопептидів з доменами SH2.

Наразі відомо близько 10 низькомолекулярних структур, здатних інгібувати домени SH2. Враховуючи таку їх незначну кількість, пошук нових

лігандів є важливим завданням, вирішення якого сприятиме створенню/виявленню нових хемотипів як елементів ефективних лікарських засобів для профілактики хвороб, пов'язаних з порушеннями функціонування доменів SH2.

Застосування комп'ютерних методів для пошуку низькомолекулярних лігандів на основі визначених тривимірних структур передбачає відбір сету потенційно активних лігандів щодо тієї чи іншої мішені. У роботі застосовано комплекс методів комп'ютерного моделювання для створення універсальної цільової бібліотеки потенційних інгібіторів доменів SH2. Тобто, отриманий сет речовин можна застосувати для пошуку активних речовин для будь-якого домену SH2 під час біологічного тестування. Активність і специфічність отриманих активних сполук з такого сету буде підвищуватись шляхом їх модифікування відносно функціональних особливостей конкретного домену SH2. Нарешті, одержані результати комп'ютерного моделювання потребують перевірки на різних біологічних моделях.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась в межах наукової тематики ННЦ «Інститут біології» за номером держреєстрації 0111U004648 «Механізми реалізації адаптаційно-компенсаторних реакцій організму за умов розвитку різних патологій» (№ 11 БФ-036-01).

Мета і завдання дослідження. Мета роботи полягає у створенні цільової бібліотеки потенційних лігандів доменів SH2 та їх перевірці на біологічних моделях.

Для досягнення цієї мети були поставлені такі наукові завдання:

1. використовуючи наявні бази даних структур доменів SH2 та активних речовин щодо них, провести розподіл цих доменів на групи/кластери, виділити шаблонні структури та відібрати репрезентативний сет речовин з наявного пулу низькомолекулярних потенційних лігандів;
2. оцінити геометричні характеристики відібраних доменів SH2 та порівняти їх;
3. виділити ключові точки зв'язування доменів SH2 з нативними лігандами;
4. створити цільову бібліотеку потенційних лігандів доменів SH2;
5. провести експериментальну перевірку отриманих сполук щодо їх впливу на проліферацію пухлинних ліній клітин K562, U937, A431 і LNCap та кіназну активність Btk та Syk.

Об'єкт дослідження – домени SH2, низькомолекулярні ліганди.

Предмет дослідження – фізико-хімічні властивості кишені зв'язування доменів SH2, структурні особливості активних речовин щодо доменів SH2, нековалентна взаємодія комплексу “домен-ліганд”.

Методи дослідження – попарне та множинне порівняння амінокислотних послідовностей; метод аналізу якості вирівнювання

амінокислотних послідовностей (оцінка консервативності та ідентичності); метод k-середніх; метод геометричного аналізу (значення величини Rmsd між структурами, які порівнюються, та доступність взаємодіючих поверхонь); чотирирівнірна енергетична мінімізація структур; гнучкий докінг; побудова прямих/структурно-залежних та непрямих/ліганд-залежних фармакофорних моделей; метод GTM (генерування топографічних карт); оцінка впливу отриманих сполук на фосфорилування протеїну STAT3, проліферацію пухлинних клітин шляхом детектуванням флуоресценції за використання резазурину та кіназну активність Btk та Syk шляхом детектування люмінесценції АДФ.

Наукова новизна отриманих результатів.

1. Усі наявні структури доменів SH2 поділено на групи за одновимірною та тривимірною подібністю.
2. Виділено основні структурні частини доменів SH2, які беруть участь у зв'язуванні з лігандом; серед них виявлено найбільш та найменш гнучкі.
3. Запропоновано оригінальний алгоритм застосування методів комп'ютерного моделювання для створення цільової бібліотеки потенційних лігандів доменів SH2.
4. Побудовано фармакофорні моделі доменів SH2.
5. Одержано нові хемотипи, зокрема бензолсульфонамід та піримідин, які потенційно можуть виступати заміниками рТур - частини пептидних лігандів при зв'язуванні з доменами SH2.
6. Встановлено речовини, які впливають на проліферацію пухлинних клітин ліній K562, U937, A431 та LNCap, проявляють інгібуючий ефект щодо домену SH2 протеїну STAT3 і здатні специфічно інгібувати кіназну активність Btk та Syk.

Практичне значення отриманих результатів.

1. Запропонований алгоритм відбору потенційних низькомолекулярних лігандів для доменів SH2 може бути ефективним для різних класів протеїнів.
2. Отримана бібліотека потенційно активних речовин може бути застосована для пошуку лігандів відносно будь-якого домену SH2 у біологічному тесті.
3. Знайдені біологічно активні сполуки можуть бути використані для розробки нових лікарських засобів.

Особистий внесок здобувача. Основна частина дисертації є самостійною науковою працею, у якій висвітлені власні ідеї/підходи щодо застосування різноманітних методів комп'ютерного моделювання для вирішення поставлених завдань. Автор самостійно опрацював літературні та експериментальні дані. Планування та обговорення усіх етапів дослідження проведені за участі наукового керівника, співробітників хімБіоЦентру Київського національного університету імені Тараса Шевченка та професора А. Варнека (Франція, Університет м. Страсбург). Аспірант виконав: розподіл доменів SH2 на групи і кластери; провів аналіз кристалографічних даних

доменів; проаналізував відомі активні речовини стосовно доменів; провів молекулярний докінг; побудував фармакофорні моделі; виконав віртуальний і частково біологічний скринінги; побудував GTM моделі; оцінив енергетичні характеристики зв'язування при утворенні комплексів “ліганд-домен”; самостійно встановлював та налаштовував програмне і апаратне забезпечення на усіх етапах розрахунків; активно обговорював отримані результати, займався написанням наукових статей і висвітленням результатів на конференціях.

Апробація результатів дисертації. Результати досліджень були представлені і доповідались на VII міжнародній науково-технічній конференції «БФФХ-2012» (23-27 квітня 2012 р., м. Севастополь, Україна); 2nd Ukrainian-French School-Seminar «Carbon Nanomaterials: Structure and Properties» (16-20 September 2013, Beregove, Ukraine); XI міжнародній науковій конференції студентів та аспірантів “Молодь і поступ в біології” (Львів, 20-23 квітня 2015р.); V Ukrainian–German Symposium on Physics and Chemistry of Nanostructures and on Nanobiotechnology (21–25 September 2015, Kyiv, Ukraine); International Conference for Young Scientists (21–25 September 2015, Kyiv, Ukraine), а також на наукових семінарах кафедри біофізики та ННЦ “Інститут біології” Київського національного університету імені Тараса Шевченка.

Публікації. За матеріалами дисертації у наукових виданнях опубліковано 11 праць: 6 статей [Гурмач В.В. та ін., 2012, Гурмач В. В. та ін., 2012, Гурмач В.В. та ін., 2015, Hurmach V. V. та ін., 2014, Hurmach V. V. та ін., 2015, Hurmach V. V. та ін., 2015] (одна з яких є оглядовою) у вітчизняних наукових фахових виданнях, з них 3 статті у журналах, включених до наукометричної бази даних Scopus; 5 тез наукових доповідей на міжнародних конференціях [Гурмач В. В. та ін., 2012, Hurmach V. V. та ін., 2013, Hurmach V. V. та ін., 2015, Hurmach V. V. та ін., 2015, Hurmach V. V. та ін., 2015].

Структура та обсяг дисертації. Дисертаційна робота викладена на 135 сторінках тексту, складається зі вступу, огляду літератури, опису матеріалів та методів дослідження, результатів дослідження та їх обговорення, узагальнення результатів дослідження, висновків і списку літератури, який включає 209 джерела, з них 205 англомовні. Дисертаційна робота містить 19 таблиць, 22 рисунки.

ОСНОВНИЙ ЗМІСТ РОБОТИ

Матеріали та методи досліджень

Усі розрахунки проводили на обчислювальному кластері Київського національного університету імені Тараса Шевченка [Boyko Y.V. та ін., 2003], використовували базу хімічних речовин підприємства Enamine, яка містить понад 1 млн. [Churpina A. та ін., 2010]. Основні підготовчі етапи розрахунків

виконані за допомогою оригінальних програм та bash-скриптів [Daniel R. та ін., 2013], написаних на мові C++ та java [Naugler D. та ін., 2007].

Вихідні 3D структури доменів SH2 взяті з баз PDB [Berman H.M. та ін., 2000] та Pfam [Punta M. та ін., 2012]. Загалом відібрали 219 структур доменів SH2, з яких 66 одержані методом ЯМР у розчині і 153 – методом рентгеноструктурного аналізу. Також врахували, що структурні файли містили декілька різних копій одного і того ж домену. Як наслідок, ЯМР - структури розбили на окремі моделі, а структури, отримані методом рентгеноструктурного аналізу, на окремі протеїнові ланцюги. Загалом отримали 1129 окремих структур доменів SH2.

Одержані структури порівняли у програмному пакеті ClustalX [Thompson J.D. та ін., 1997]. Зокрема, застосували множинне порівнювання амінокислотних послідовностей (MSA) [Ramu S. та ін., 2003] за прогресивним методом [Hogeweg P. та ін., 1984]. Порівняння відбувалося у три етапи: (1) попарне порівняння [Mehmet K. та ін., 2006]; (2) побудова дендрограми; (3) множинне вирівнювання послідовностей. Також оцінили загальну площу кожної амінокислоти у сайті зв'язування домену. Це дало змогу визначити ключові амінокислоти сайту зв'язування доменів SH2: так, якщо амінокислота має площу поверхні більшу, ніж 5 \AA^2 , і перебуває на відстані 5 \AA , або менше до будь-якого атома зв'язаного ліганду, вона визначається як амінокислота зі значним внеском у формування сайту зв'язування; з іншого боку, амінокислоти, які мають меншу площу поверхні, або знаходяться на більшій відстані від будь-якого атома зв'язаного ліганду, визначаються як допоміжні структури без значного внеску у формування сайту зв'язування.

Для отриманих репрезентативних структур доменів SH2 провели конформаційний пошук методом Low Mode Search (коротка молекулярна динаміка, 0.1 пс).

На противагу поділу за подібністю амінокислотних послідовностей, провели кластеризацію дескрипторів (виділяли ключові амінокислоти при їх порівнянні) сайтів зв'язування на основі заданої метрики методом k-середніх [Vattani A. та ін., 2011].

Як пул з набором потечійних лігандів, використали базу хімічних речовин підприємства Enamine [Chuprina A. та ін., 2010]. Провели молекулярний докінг у двох варіантах: у першому базу хімічних речовин скоротили до 50 (докування проводили за допомогою програмного пакету Flo+ [Mcsmartin C. та ін., 1997]); в іншому варіанті для проведення докінгу використовували усю базу речовин із застосуванням пакетів MOE [Corbeil C. R. та ін., 2012; Chang D. T. та ін., 2005] та Flo+ [Mcsmartin C. та ін., 1997].

На наступному етапі, з урахуванням усіх речовин з відомою активністю щодо доменів SH2 та наявних PDB комплексів “домен SH2 – нативний ліганд”, з використанням програмного пакету LigandScout [Gerhard W. та ін., 2005] побудували одну структурно-залежну та вісім ліганд-залежних

фармакофорних моделей. Для валідації отриманих моделей використали гос-аналіз [Braga C. та ін., 2013]. Зазначимо, що спочатку за допомогою інтернет-ресурсу «DUDE decoy generation» згенерували 4200 декоїв, які потім були застосовані для валідації. Після цього на виділених фармакофорних моделях провели віртуальний скринінг.

Результати, одержані на попередніх етапах дослідження (сет активних і неактивних речовин, результати віртуального скринінгу), були використані для побудови моделей GTM за алгоритмом, запропонованим у роботах [Hélène A. G. та ін., 2013; Kireeva N. та ін., 2012].

Аналіз впливу відібраних для дослідження речовин на індуковане фосфорилування STAT3 проводили методом імуноблот-аналізу [Braga C. та ін., 2013] з використанням клітинної лінії L540 (лімфогранулематоз), для якої характерний високий рівень експресії фосфорильованого STAT3. Клітини культивували за температури 37° у зволоженому середовищі у змішаній атмосфері повітря - 95% O₂ і 5% CO₂. Концентрації речовин склали 30 μM, а час впливу – 24 год.

Оцінку проліферативної активності речовин проводили методом флуоресцентного детектування проліферації клітин з використанням резазурину. В такому експерименті кількість клітин після інкубації (з додаванням або без додавання до розчину потенційного ліганду) визначали за інтенсивністю флуоресценції резозурину, що є продуктом відновлення резазурину в клітинах. Як контроль, використовували доксорубіцин, як відомий інгібітор проліферації клітин, механізм дії якого полягає в інтеркаляції ДНК та інгібуванні біосинтезу. Експеримент проводили у 384-коміркових (40 μл/комірка) платах у дублікатах. Результуюче значення інгібування визначали як середнє з трьох повторів. Дослідження проводили на чотирьох лініях пухлинних клітин: A431 (плоскоклітинна карцинома), K562 (хронічний мієлоцитарний лейкоз), U937 (гістоцитна лімфома) та LNCap (аденокарцинома простати). Їх інкубували у таких середовищах: K562 - 5×10^5 клітин/мл у середовищі RPMI-1640 з додаванням пеніциліну/стрептоміцину та L-глутаміну; U937 - 7×10^5 клітин/мл у середовищі RPMI-1640 з додаванням пеніциліну/стрептоміцину та L-глутаміну; A431 та LNCap - 5×10^5 клітин/мл у середовищі DMEM пеніциліну/стрептоміцину та L-глутаміну.

Для визначення життєздатності вихідних тест – клітин кількість резазурину доводили до концентрації 50 мкM та інкубували 2 год у зволоженої атмосфері за 37°C та 5% CO₂. Кількість резазурину (пропорційна числу життєздатних клітин) оцінювали шляхом вимірювання флуоресценції при E_x - 490 нм, E_m - 540 нм з використанням FLIPR® (Fluorometric Imaging Plate Reader).

Для сполук, обраних за результатами клітинного скринінгу, визначали їх вплив на кіназну активність Syk та Btk (Homo sapiens). Ці кінази містять 2 та 1 домени SH2, відповідно. Активність Syk та Btk кіназ визначали з

використанням наборів Btk (cat. V2941) та Syk (cat. V3801) kinases (Promega corporation; Madison, WI, USA) згідно стандартного протоколу. Дослідження починали з додавання 2 μL АТФ маточного розчину до кінцевої концентрації 50 μM . Після 1,5 год збовтування (300 rpm) за температури 27°C додавали ADP-Glo реагент (7 μL). Для оцінки люмінесценції через 40 хв додавали люциферазу (14 μL). Після 30 хв інкубації вимірювали люмінесценцію за допомогою BMG Polarstar Omega рідера. Активність кіназ є пропорційною інтенсивності люмінесценції. Як позитивний контроль, використали сигнал з комірок, у яких досліджувані реагенти інкубували без субстрату (нульова кіназна активність). Як негативний контроль, використали сигнал з комірок, куди не додавали потенційний ліганд. Експеримент проводили у квадруплікатах, тобто кожна сполука була додана до чотирьох комірок плати, сигнал по яких усереднювався. Результати виражали у відсотках інгібування кіназної активності.

Результати досліджень та їх обговорення

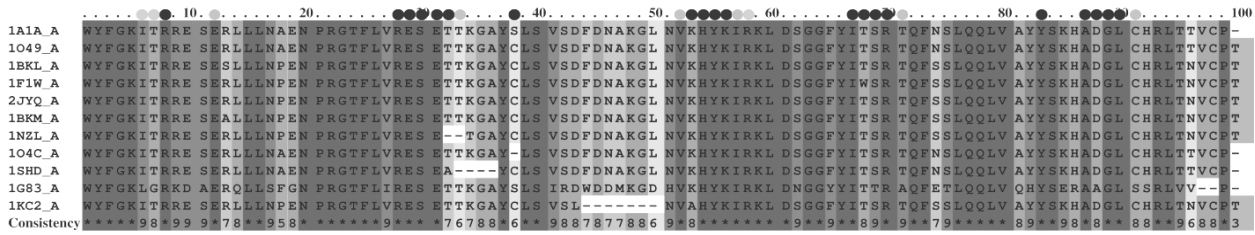
Відбір репрезентативних структур доменів SH2. Поетапно застосували методи структурного аналізу (геометричний аналіз, множинне вирівнювання амінокислотних послідовностей). Усі наявні домени SH2 поділили на 1129 окремих копій, з яких виділили 56 унікальних структур, які потім поділили на шість груп з урахуванням усієї довжини амінокислотних послідовностей доменів (рис. 1.) У пакетах Chimera, PRALINE та Jalview порівняли одержані структури і групи. Оцінка якості цих груп виявила високий та середній рівні консервативності (група 1 – 92/91%, група 2 – 81/87%, група 3 – 56/62%, група 4 – 37/51%, група 5 – 50/68%, група 6 – 51/61%) та незначні коливання значень Rmsd (група 1 – 0.8 Å, група 2 – 1.16 Å, група 3 – 1.6 Å, група 4 – 1.9 Å, група 5 – 2 Å, група 6 – 1.4 Å).

У всіх випадках сайт зв'язування виявився консервативним: група 1 – 94/93%, група 2 – 89/96%, група 3 – 59/67%, група 4 – 49/60%, група 5 – 53/70%, група 6 – 53/59%. Групи 1 і 2 є найбільш консервативними, зростання їх ідентичності (якщо порівнювати з усіма доменами) складає 61% і 54%, а подібності - 55.1% і 35.2%, відповідно. Групи 3-6 є менш консервативними порівняно з першими двома: зростання ідентичності (порівняно з усіма доменами) становить 19%, 14%, 8% і 14%, а подібності - 14,5%, 12,2%, 15,5 і 11,63%, відповідно, що, ймовірно, і призводить до значного збільшення площі активної поверхні доменів.

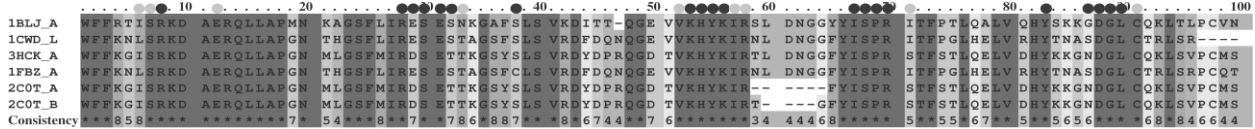
У середньому кишеня зв'язування займає $\frac{1}{4}$ від поверхні домену. У кожному випадку вона є плоскою і містить 20-24 амінокислот, які формують ключові точки кишені зв'язування. Хоча, варто відзначити, групи 3, 4 і 6 містять ще одну додаткову точку.

Основні частини кишені зв'язування доменів SH2 формуються з βB - і βG -складок та αB -спіралі (рис. 1, 2; табл. 1). Загалом її можна поділити на такі ділянки: рТуг-зв'язуюча, яка формується послідовністю ITR (ділянка 1),

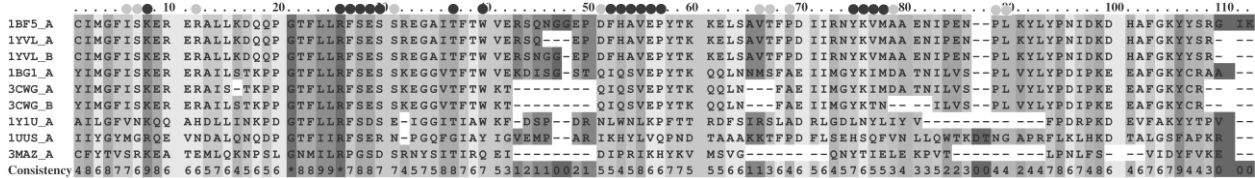
Група 1



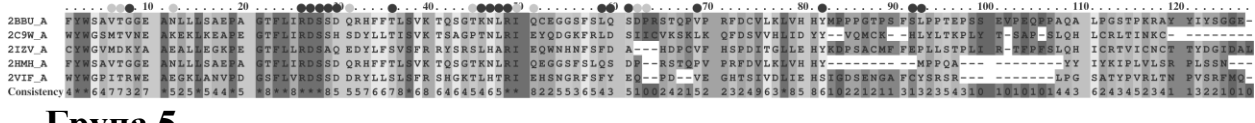
Група 2



Група 3



Група 4



Група 5



Група 6



Рис. 1. Множинне вирівнювання послідовностей доменів SH2. Основні амінокислотні залишки сайту зв'язування позначені чорними крапками (утворюють основну частину сайту зв'язування), сірими позначено амінокислотні залишки, внесок яких в утворення сайту зв'язування є незначним.

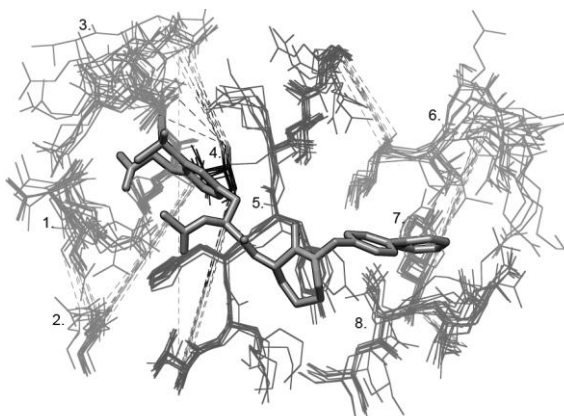


Рис. 2. Сайт зв'язування групи 1 з нативним лігандом і розподіл його ділянок: ділянка 1 (ITR), ділянка 2 (E), ділянка 3 (FLVRESETT), ділянка 4 (C), ділянка 5 (VKHYKIR), ділянка 6 (ITSRT), ділянка 7 (Y) і ділянка 8 (ADGLC).

SH2 domain of proto-oncogene tyrosine-protein kinase, Homo sapiens); 3IN7 (crystal structure of the Grb2 SH2 domain in complex with a Cyclopropyl-constrained Ac-pY-Q-N-NH₂ Tripeptide Mimic, Homo sapiens), 2K7A (complex between the SH3 and SH2 domain of interleukin-2 tyrosine kinase, Mus musculus), 2JYQ (NMR structure of the apo v-Src SH2 domain, Rous sarcoma virus), 2GE9 (solution structures of the SH2 domain of Bruton's Tyrosine Kinase, Homo sapiens), 2FCI (Syk tyrosine kinase, Bos taurus), 1UUS (structure of an activated D. STAT in its DNA-unbound form, Dictyostelium discoideum) і 1O49 (crystal structure of SH2 proto-oncogene tyrosine protein kinase SRC, Homo sapiens).

Конформаційний аналіз виявив рухливість доменів SH2 в межах (1.5–5.5) Å. Так, рТур–зв'язуюча та гідрофобна частини є доволі гнучкими, а центральна – жорсткою. Для наведених вище репрезентативних структур згенерували 39 можливих конформацій за мінімального значення Rmsd. Загальна рухливість сайту зв'язування є меншою ніж усього домену, що є очікуваним, оскільки сайт зв'язування є більш консервативним, аніж усі інші структурні частини домену. Отже, його середня рухливість для різних структур становить: 1O49 - 2.26 Å, 1UUS - 1.27 Å, 2FCI - 2.27 Å, 2GE9 - 2.29 Å, 2JYQ - 2.12 Å, 2K7A - 2.38 Å, 2KK6 - 2.28 Å і 3IN7 - 1.74 Å. Звичайно, кожна частина сайту зв'язування дає свій внесок у гнучкість домену. Так, рТур–зв'язуюча частина (утворюється AA-петлею, αA-спіраллю, βB-складкою,

FLVRESETT (ділянка 3) та C (ділянка 4) (причому лише послідовність RESETT взаємодіє з лігандом); центральна частина - ділянки 2 (E) і 5 (VKHYKIR) (βD-складка та αA-спіраль); гідрофобна частина - ділянка 6 (ITSRT), ділянка 7 (Y) і ділянка 8 (ADGLC) (βG-складка, αB- та αA-спіралі). Виявили чотири висококонсервативні позиції (>70%), якщо брати до уваги всі домени SH2: позиція 3 ділянки 1, позиції 1 і 3 ділянки 3 та позиція 3 ділянки 5. Також знайшли подібні позиції, які не є частиною будь-якого із сайтів зв'язування з вище описаних, наприклад, T (позиція 37 - група 3), L (позиція 69 - група 4) та Q (позиція 40 - група 6).

Загалом отримали вісім найбільш репрезентативних сайти зв'язування доменів SH2: 2KK6 (solution structure of

Таблиця 1

Розташування амінокислот сайту зв'язування доменів SH2. Параметри окремих груп доменів SH2 та їх сайту зв'язування показані як Рос.-сайт зв'язування і Р.-весь протеїн. Ділянки зв'язування показані в [...]. Значення консервативності наведені в одній колонці ($K = 21/K = 7$)

Група та шаблонний PDB	Амінокислотне оточення сайту зв'язування	Ідентичність сіквенсів, %		Середнє значення Rmsd, Å		Середня площа поверхні, Å ²		Значення консервативності	
		Рос.	Р.	Рос.	Р	Рос.	Р	Рос.	Р
Group 1 1A1A	[1]. IWR. [2]. Q. [3]. RQSQQW. [4]. S. [5]. VKHYKIR. [6]. IWSRW. [7]. Y. [8]. ADGLC.	0.92	0.91	0.72	0.79	1268	5876	0.94/ 0.93	0.92/ 0.91
Group 2 1BLJ	[1]. ISR. [2]. S. [3]. RQSQSN. [4]. S. [5]. VKHYKIR. [6]. ISPRI. [7]. Y. [8]. GDGLC.	0.85	0.78	1.14	1.32	1491	5856	0.89/ 0.96	0.81/ 0.87
Group 3 1BF5	[1]. ISK. [2]. R. [3]. RFSQSS. [4]. W. [5]. DFHAVQP. [6]. VWP. [7]. YKVMA. [8]. PL. [9]. W.	0.5	0.47	1.53	1.65	1552	6945	0.59/ 0.67	0.56/ 0.62
Group 4 2C9W	[1]. MWV. [2]. KRDSHS. [4]. W. [5]. PWNLRIQ. [6]. LDSII. [7]. Y. [8]. HL. [9]. L.	0.45	0.32	1.82	1.99	1355	7161	0.49/ 0.60	0.37/ 0.51
Group 5 2DMO	[1]. IWR. [2]. Q. [3]. RDSRHL. [4]. W. [5]. IKHYQKI. [6]. VAQRH. [7]. H. [8]. AQL.	0.39	0.39	1.81	2.00	1645	6499	0.53/ 0.70	0.50/ 0.68
Group 6 1WQU	[1]. IPR. [2]. A. [3]. RQSQG. [4]. V. [5]. PRHFIIQ. [6]. LQG. [7]. L. [8]. QPLWKS. [9]. Q.	0.45	0.45	1.27	1.45	1639	5429	0.53/ 0.59	0.51/ 0.61

BC-петлею і β C-складкою) і гідрофобна кишеня (β B-складка, BC-петля, β C-складка і “DE-петля, β E-складка, EF-петля і FB-петля”) є високорухливими. Рухливість рТуг–зв’язуючої частини лежить в межах (1.1–2.5) Å. Відзначимо також, що практично усі вторинні структури дають значний внесок у рухливість рТуг–зв’язуючого сайту: наприклад, AA-петля - 2.8 Å, α A-спіраль - 2.7 Å, β C-складка - 2.7 Å, BC-петля - 2.8 Å (за винятком β B-складки - 1.6 Å). Середня рухливість гідрофобної частини кишені становить: 1O49 – 2.46 Å, 1UUS – 1.63 Å, 2FCI – 2.3 Å, 2GE9 – 2.6 Å, 2JYQ – 2.4 Å, 2K7A – 2.5 Å, 2KK6 – 2.5 Å і 3IN7 – 1.7 Å. Для різних центроїдів внески у рухливість (залежно від вторинної структури) є різними, наприклад: для 1O49, 2FCI і 2GE9 BG-петля є високорухливою (2.7 Å, 2.3 Å і 3 Å, відповідно), а для 1UUS і 3IN7 її рухливість становить 1.3 Å і 1.7 Å, відповідно. Найбільш стабільною виявилась центральна частина кишені зв’язування. Лише у двох випадках (для 2FCI і 2GE9) середня рухливість перевищує 2 Å. У всіх інших випадках рухливість є меншою: для 1O49 – 1.3 Å, 1UUS – 1 Å, 2JYQ – 1.8 Å, 2K7A – 2 Å, 2KK6 – 1.8 Å і 3IN7 – 1.6 Å. Практично у всіх випадках β D-складка і α A-спіраль формують рТуг- зв’язуючу частину. Але і тут існують винятки:

Таблиця 2

Результати конформаційного пошуку репрезентативних структур доменів SH2

PDB	Rmsd, Å	Загальна кількість отриманих конформацій	Rmsd, Å (порівняно з нативним доменом)
1O49	0.25	3	3.2
			2.6
			5.17
	0.5	1	1.34
	0.75	1	1.44
1UUS	0.25	1	1.09
	0.5	2	1.15
			3.37
0.75	1	2.8	
3IN7	0.25	2	0.94
			3.74
	0.5	1	0.95
	0.75	1	2.39
2FCI	0.25	2	2.78
			2.12
	0.5	2	3.14
			2.07
0.75	1	2.61	
2GE9	0.25	1	2.67
	0.5	2	2.49
			2.26
	0.75	2	2.17
			3.10
2JYQ	0.25	1	2.3
	0.5	2	2.54
			2.33
	0.75	3	2.79
			2.6
			2.62
2K7A	0.25	1	2.75
	0.5	2	2.11
			3.21
	0.75	3	2.87
			2.95
			2.03
2KK6	0.25	1	2.31
	0.5	1	2.21
	0.75	2	2.32
			3.14

центральну частину 1UUS також формують BC-складка (TYR 625) і “DE-петля, BE-складка, EF-петля і FB-петля”, амінокислоти VAL 639 і GLN 640;

2FC1 структуру - CD-петля; 2K7A структуру - BC-складка і “DE-петля, β E-складка, EF-петля і FB-петля”.

Пошук потенційних лігандів для доменів SH2. Редокінг виявив задовільні як геометричні, так і енергетичні можливості відтворення застосованих методів (значення Rmsd знаходиться в межах 2 Å).

За допомогою молекулярного докінгу виявили чотири основні моделі зв'язування типу “домен-ліганд”

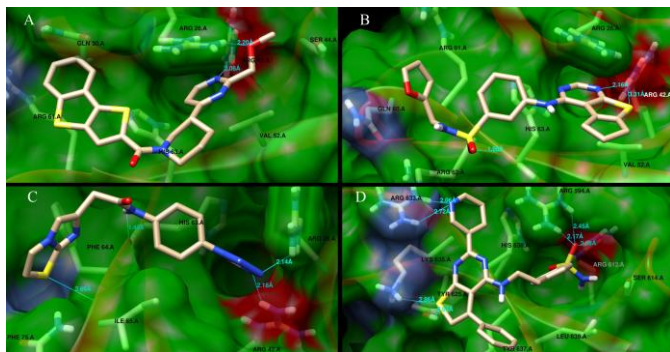


Рис. 3. Варіанти зв'язування доменів SH2 на прикладі центрів 2KK6 (A, B, C) та 1UUS (D) (стрічкова та поверхнева діаграми).

(рис. 3) та декілька хімічних груп, які здатні імітувати фосфотирозин (рис. 4):

A. водневі зв'язки утворюються з Arg 26 (також має місце катіон- π взаємодія) і Arg 46. Спостерігається стерична взаємодія з Gln 30 і Val 52 не основної гідрофобної кишені та центральним His 63;

B. стабілізація відбувається завдяки взаємодіям з Arg 26 (катіон- π) і Arg 42 (воднева). Важливим є утворення водневого зв'язку між His 63 та арилсульфомідною групою, що сприяє міцному зв'язуванню ліганда з доменом;

C. утворення міцних водневих зв'язків з Arg 42 та Arg 26 і NH групи потенційного ліганда з His 63 (центр кишені зв'язування). Наявність стеричних взаємодій з Gln 67, Ile 65, His 63, Phe 75 та Phe 64. Структури, відібрані в рамках цієї моделі, виявили значний

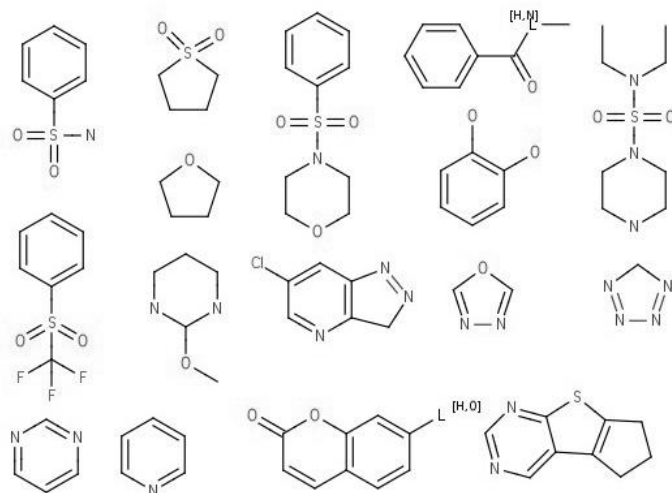


Рис. 4. Фрагменти структур, які виступали як імітатори фосфотирозину.

інгібуючий ефект;

D. модель принципово відрізняється за механізмом зв'язування. Фосфотирозин зв'язуюча кишеня щільно заповнюється, утворюються міцні водневі зв'язки з Arg 594 і Arg 612. Arg 633 утворює другий кластер водневих зв'язків (потенційно інший фосфотирозин-зв'язуючий сайт). Спостерігається значна стекінг взаємодія з центральною частиною кишені. Речовини цієї моделі також виявили значний інгібуючий ефект.

Відібрані моделі відображають основні варіанти зв'язування, які виявили після фільтрування. Як бачимо з табл. 3, загальна енергія зв'язування (FreE) є мінімальною для моделі D, але у цьому випадку спостерігається значне внутрішнє напруження ліганду (Intl), незначна контактна енергія (Cntc) і відносно високе значення стеричних зіткнень (Bump). У моделях A і B спостерігається практично однакова Cntc та FreE, а у моделях A і C – Intl та Bump.

Таблиця 3

Енергетичні параметри зв'язування “домен-ліганд” (кДж/моль)

модель	Cntc	Hbond	Intl	Bump	FreE
A	-52.8	-1.9	3.8	2.5	-16.0
B	-58.5	-2.7	6.6	4.6	-16.6
C	-47.9	-2.9	3.3	2.9	-19.5
D	-71.6	-4.7	8.2	4.1	-23.5

Такі відмінності пов'язані зі складностями зв'язування компактного фосфотирозинового сайту з великим поліциклом. З цього випливає, наприклад, що відтворення моделі B є менш ймовірним, аніж A. Згідно з наведених результатів, модель C повинна бути найбільш відтворюваною. У цьому випадку спостерігається найбільша Cntc, практично мінімальні Bump, FreE та Intl.

У клітинному біологічному тесті виявили 56 лігандів з моделі C, 14 з D та 13 з A. Модель B не підтвердилась у жодному з випадків (можливо це пов'язано зі складностями зв'язування масивного гетероциклічного ядра з малим фосфотирозиним сайтом зв'язування домену).

Загалом отримали три бібліотеки потенційних лігандів для доменів SH2. Основна з них склала 10463 сполук (як результат молекулярного докінгу): перекривання між різними центроїдами становило 705 речовин.

Усього побудували дев'ять фармакофорних моделей: одна структурно-залежна і вісім ліганд-залежних. У всіх випадках зв'язування характеризується такими ключовими моментами (рис. 5, табл. 4): (1) зв'язування з рТуг/карбоною кислотою, велика кількість донорів Н-зв'язку (Arg, Lys); (2) ароматична/гідрофобна частина ліганду, Ван-дер-ваальсова взаємодія здебільшого між Arg і Lys; (3) Ніс утворює водневий зв'язок з групою NH пептидного ліганду; (4) взаємодія у гідрофобній кишені, яка знаходиться за центром.

Ліганд-залежні моделі № 1, 3, 6, 8 є найкращими (AUC 100%), № 4, 5 і 7 трохи гірші (AUC 75%). Моделі № 1, 2, 6 і 8 найкраще відбирають активні речовини і відсіюють декої; модель №3 відбирає речовини, надані тренінг-сетом, але відкидає декої; протилежна ситуація спостерігається у моделі №5. Так, всі отримані моделі доповнюють одна одну, частина відбирає усі активні

речовини, інші відкидають більшість декоїв. Моделі № 3 і 5 не відповідають геометрії сайту зв'язування (причиною може бути здатність набувати інших конформацій молекулами, з яких сформовані моделі).

Після віртуального скринінгу одержали 1816 активних речовин. Ці речовини у подальшому аналізували за допомогою методу GTM, в рамках якого розглянуто 78 активних і 43 неактивних речовин, які поділили на 4155 дескрипторів. Застосували такі параметри моделі GTM: k (number R of RBF) (25), m (the grid resolution) (5), w (the RBF width) (0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2), l (the weight regularization coefficient) (0,01, 0,1, 1, 10, 100).

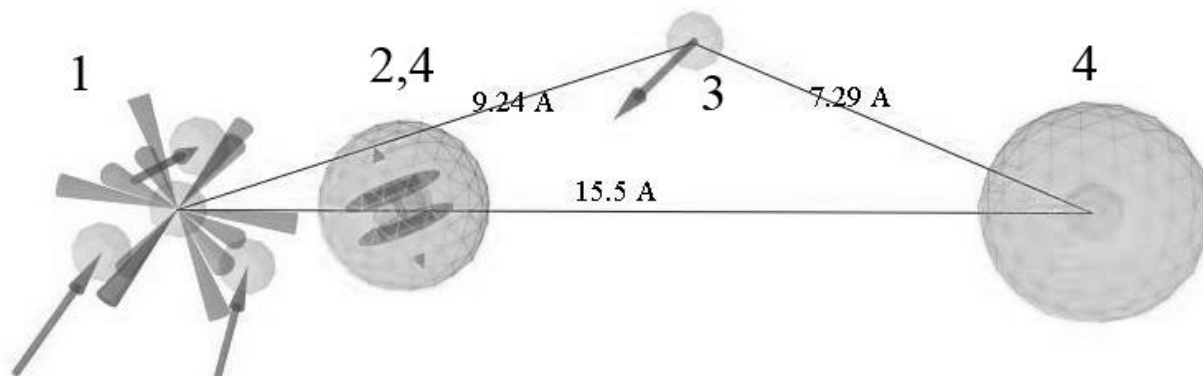
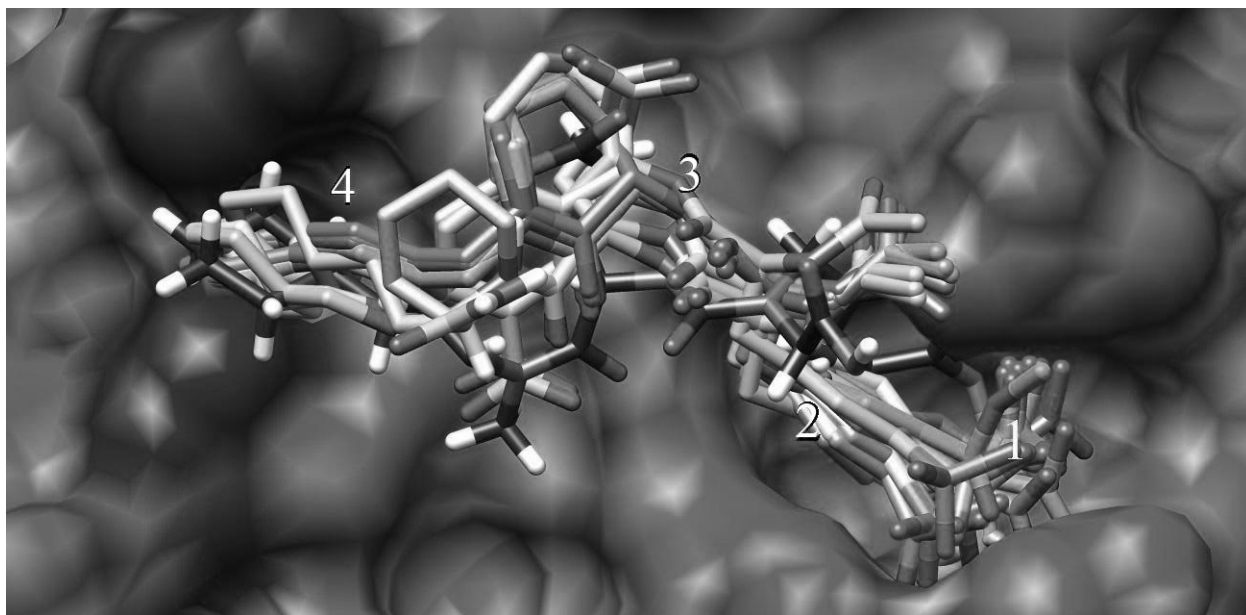


Рис. 5. Накладання усіх наявних кристалів доменів SH2 з лігандами (верхня частина рисунку). Також показано чотири ключові точки зв'язування (1, 2, 3 і 4), на основі яких була побудована загальна фармакофорна модель (на моделі (нижня частина рисунка): 1 - акцептор водневого зв'язку, 2 - ароматична частина ліганду, 3 - донор водневого зв'язку, 4 - гідрофобна частина ліганду).

Таблиця 4

Моделі, побудовані на основі інгібіторів доменів SH2: результати їх валідації і віртуального скринінгу

№	Фармакофорна модель (Позначення 1 – 4 відповідають позначенням фармакофорної моделі на рис. 5)	Репрезентативна речовина	AUC (area under ROC curve)	EF (enri- chment factor)	THR, % true hit rate	FHR, % false hit rate	хіти	
							Тест сет	Декої
1			1.00	559	100	0.12	90	5
2			0.92	318	83	0.14	159	6
3			1.00	54.5	100	1.79	400	75
4			0.75	39.6	50	1.23	604	52
5			0.75	91.2	50	0.59	216	25
6			1.00	233	100	0.33	129	14
7			0.75	110	100	0.35	334	15
8			1.00	209	100	0.41	522	17

Одержали чотири найкращі варіанти розрахунків, які сформовані сімома кластерами речовин (сульфіди, карбонові кислоти, фосфати і фосфанати, паразаміщені арілазоли та інші) (рис. 6): середнє значення $T_c=0.22$ (Tanimoto coefficient) - між кластерами і $T_c=0.92$ – всередині кожного кластеру. Водночас спостерігали часткову подібність з неактивними речовинами: максимальне значення - $T_c=0.56$, середнє значення - $T_c=0.26$ і мінімальне - $T_c=0.12$. Використовуючи одержані моделі, мало місце передбачення за кластерами 2, 5, 6 і частково 7. Результуюча бібліотека потенційних лігандів для доменів SH2 склала 1607 речовин.

Біологічне тестування отриманих структур. Потенційні ліганди для домену SH2 протеїну STAT3 перевірили на L540 лінії клітин. Речовини тестували за концентрації 30 мкмоль/л упродовж 24 год. Найбільш вдалою виявилася модель № 4 (хоча за теоретичними розрахунками вона гірша за

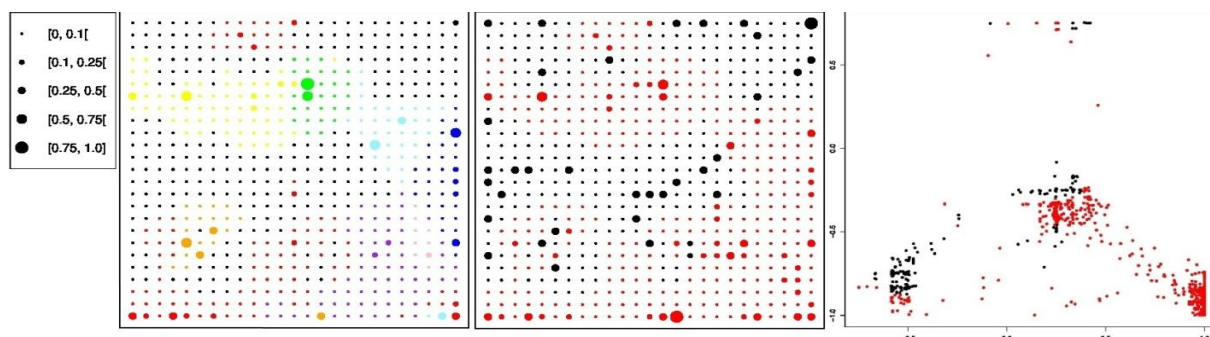


Рис. 6. GTM візуалізація одержаної моделі. Карта зліва - GTM модель з урахуванням розбиття усіх активних речовин на різні кластери: чорний колір – кластер 1 (неактивні речовини (не враховували при наведені вищедосліджуваних кластерів)); червоний колір – кластер 2 (сінгелтоно, речовини, які не піддалися кластеризації); синій колір – кластер 3 (дигідрокси/метокси-феніланіні); зелений колір – кластер 4 (тріазолпірадозіни-6-пірідіні); помаранчевий колір – кластер 5 (арілсульфаміди та арілсульфони); пурпурний колір – кластер 6 (карбонові кислоти); блакитний колір – кластер 7 (фосфати і фосфанати); зелений колір – кластер 8 (паразаміщені арілазоли). Карта в центрі - GTM модель без урахування розбиття на кластери (червоний – активні, чорний – неактивні речовини). Карта справа - результат передбачення в рамках GTM моделі.

модель № 3). Найбільш активними виявилися ліганди, в яких рТуг частину замінював скефорд з атомом кисню. Можливо це пов'язано з координацією Pro 715. Активні речовини, отримані у цьому тесті, застосували для побудови непрямих/ліганд-залежних фармакофорних моделей (табл. 4). Загалом отримали 16 активних сполук.

Репрезентативний сет результуючої бібліотеки потенційних лігандів для доменів SH2 та випадкової бібліотеки протестували на лініях пухлинних клітин K562, U937, A431 і LNCap (у кожній гіперекспресується протеїн, який містить домен SH2), оцінивши їх проліферативний ефект. Загалом виявили

Кількість хітів у відібраній бібліотеці доменів SH2 та у випадковій вибірці

Лінія клітин	K562	U937	A431	LNCap	Всього
Випадкова вибірка	5	7	26	19	57
Відібрана бібліотека доменів SH2	19	3	55	6	83

83 активних речовини у відібраній бібліотеці та 57 у випадковій (табл. 5). З активних речовин за моделлю А – 13, В – не містить жодної, С – 56 і D – 14.

Інгібітори, виявлені у клітинному скринінгу, застосували у біохімічному експерименті щодо оцінки їх впливу на кіназу активність Vtk та Syk кіназ людини. Використали метод люмінесцентної детекції АДФ. Як результат, для Vtk та Syk знайшли, відповідно, 31 та 1 активні сполуки.

ВИСНОВКИ

1. Домени SH2 є висококонсервативними структурами, особливо в межах сайту зв'язування. Найбільш консервативним є рТуг-зв'язуюча частина домену, а найменш – гідрофобна частина кишені зв'язування.

2. За геометричними параметрами домени SH2 є практично ідентичними (значення Rmsd змінюється в межах 2 Å). Вони характеризуються незначною рухливістю, яка змінюється в межах 1.5–5.5 Å. Найбільш рухливою виявилась α A-спіраль, а найменш – β D-складка.

3. Отримані групи доменів SH2 характеризуються значною подібністю: в межах кожної окремої групи спостерігаються домени лише одного або подібних типів протеїнів.

4. Окрім консервативних ділянок в межах кожної окремої групи виявили чотири висококонсервативних позиції (>70%): позиція 3 ділянки 1, позиції 1 і 3 ділянки 3 та позиція 3 ділянки 5. Усі інші частини сайту зв'язування є менш консервативними (<70%). Також встановили декілька позицій, які не є унікальними для тієї чи іншої групи: T (позиція 37 - група 3), L (позиція 69 - група 4) та Q (позиція 40 - група 6).

5. Виділено ключові точки зв'язування типу “домен SH2-ліганд”: (1) зв'язування з рТуг/карбоною кислотою (характеризується великою кількістю донорів Н-зв'язку (Arg, Lys)); (2) біля рТуг-зв'язуючої частини знаходиться ароматична/гідрофобна частина ліганду; (3) центр кишені зв'язування визначається позицією O амінокислотного залишку другої амінокислоти (His); (4) наявність однієї або двох гідрофобних кишень, які знаходяться за центром активного сайту.

6. Отримали чотири основні моделі молекулярного докінгу зв'язування типу “домен SH2-ліганд”. Практично у всіх випадках ліганд

щільно заповнює фосфотирозинову кишеню зв'язування та бере участь у водневих взаємодіях з ключовими амінокислотами Arg або Lys. Крім того, присутні стеричні взаємодії між лігандом та амінокислотами Glu, Val, Pro, His і Ser. Під час біологічного тестування, модель, яка передбачає зв'язування з масивними гетероциклами не підтвердилась. Отримали три бібліотеки потенційних лігандів доменів SH2, що склали: 1015, 3018 і 10463 структур (останні дві використані у біологічному тестуванні).

7. Побудовано дев'ять фармакофорних моделей, які відображають декілька варіантів зв'язування доменів SH2 з лігандом. Валідація показала, що усі одержані фармакофори можуть бути застосовані для пошуку потенційних лігандів доменів SH2. Значення величини AUC в отриманих моделях знаходяться в діапазоні 75-100%. Одна частина моделей краще відбирає активні речовини і відсіює декої (моделі 1, 2, 3, 6, 8 визначені, як найкращі), а інша (моделі 4, 5, 7) – краще відсіює декої або відбирає активні речовини. Як наслідок, відібрано 1816 речовин, які застосували у GTM передбаченні і одержали 1607 речовин, що відповідають ключовим точкам взаємодії, застосованих під час докінгу та фармакофорного пошуку.

8. Перевірка отриманих потенційних лігандів доменів SH2 щодо їх впливу на проліферативну активність пухлинних клітин та кіназну активність Vtk та Syk кіназ людини виявила значний інгібуючий ефект (65-95%).

СПИСОК ПРАЦЬ ОПУБЛІКОВАНИХ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових фахових виданнях

1. Гурмач В. В. Метод молекулярного докінгу за участю SH2-доменів / **В. В. Гурмач**, О. М. Балинський, М. О. Платонов, П. О. Бориско, Ю. І. Прилуцький // Біотехнологія. – 2012. – Т. 5. – С. 31 – 40. (Здобувач – опрацював літературу, узагальнивши усі наявні дані щодо доменів SH2 та ком'ютерних методів їх дослідження, підготував статтю до друку).

2. Гурмач В. В. Пошук низькомолекулярних лігандів для SH2-доменів методом докінгу міжмолекулярних взаємодій / **В. В. Гурмач**, О. М. Балинський, П. О. Бориско, М. О. Платонов, Г. Х. Баєнг, Д. Б. Ковальський, Ю. І. Прилуцький // Доповіді НАН України – 2012. – Т. 8. – С. 141 – 146. (Здобувач – провів відбір мішеней та сету потенційних лігандів для проведення молекулярного докінгу, розрахунки та їх опрацювання, аналіз отриманих біологічних результатів *in vitro*, підготував статтю до друку).

3. Hurmach V. V. Design of potentially active ligands for SH2 domains by molecular modeling methods / **V. V. Hurmach**, О. М. Balinskyi, М. О. Platonov, О. М. Boyko, Р. О. Borysko, Yu. I. Prylutskyu // Biopolymers and Cell – 2014. – V. 30. – P. 321-325. (Здобувач – провів відбір мішеней та сету потенційних

лігандів для проведення молекулярного докінгу, розрахунки та їх опрацювання, підготував статтю до друку).

4. Гурмач В. В. Застосування методів молекулярного моделювання для пошуку нових біологічно активних речовин / **В. В. Гурмач**, О. М. Балинський, М. О. Платонов, О. М. Бойко, Ю. І. Прилуцький // Укр. біохім. журн. – 2015. – Т. 87. – С. 109 – 120. (Здобувач – опрацював наявну літературу щодо сучасних методів дослідження та їх валідації, встановив та налаштував усе програмне забезпечення, провів усі розрахунки, підготував статтю до друку).

5 Hurmach V. V. Comparative analysis of SH2 domain structure / **V. V. Hurmach**, М. О. Platonov, О. М. Boyko, Yu. I. Prylutskyu // *Studia Biologica* – 2015. – V. 9. – P. 5 – 14. (Здобувач – відібрав усі найбільш репрезентативні структури, розділив їх на групи, порівняв між собою за різними параметрами, підготував статтю до друку).

6. Hurmach V. V. / New conformational properties of SH2 domain binding pocket / **V. V. Hurmach**, М. О. Platonov, Yu. I. Prylutskyu // *Ukr. Biochem. J.* – 2015. – Vol. 87. – P. 133 – 140. (Здобувач – згенерував усі можливі конформації репрезентативних доменів SH2 для різних варіантів Rmsd обмеження, проаналізував рухливість доменів загалом та їх окремих частин, підготував статтю до друку).

Тези наукових доповідей

7. Пошук низькомолекулярних лігандів для SH2 доменів методом молекулярного докінгу. / В. В. Гурмач., О. М. Балинський., М. О. Платонов., П. О. Бориско., Ю. І. Прилуцький. // VII міжнародна науково-технічна конференція БФФХ-2012. Севастополь, 23-27 квітня 2012 р.: тези доп. С. 100.

8. Design and testing of potentially active ligands for SH2-domains by molecular modelling methods. / V. V. Hurmach, А. М. Balynskyi, М. О. Platonov, А. N. Boyko., Р. О. Borysko, Yu. I. Prylutskyu. // 2nd Ukrainian-French School-Seminar “Carbon Nanomaterials: Structure and Properties” 16-20 September, Beregove, Ukraine, 2013 p.: тези доп. С. 79.

9. SH2 domain structure analysis. / V. V. Hurmach // XI міжнародна наукова конференція студентів та аспірантів “Молодь і поступ в біології” Львів, 20-23 квітня 2015 р.: тези доп. С. 12.

10. Computational study of SH2 domain. / V. V. Hurmach, А. А. Varnek, Yu. I. Prylutskyu. // Ukrainian–German Symposium on Physics and Chemistry of Nanostructures and on Nanobiotechnology, Kyiv, Ukraine, 21–25 September, 2015 p.: тези доп.: С. 262.

11. Conformational analysis of SH2 domain. / V. V. Hurmach, А. А. Varnek, О. М. Platonov, Yu. I. Prylutskyu. // International Conference for Young Scientists CYS-2015, Kyiv, Ukraine, September 21–25, 2015 p.: тези доп. С. 93.

АНОТАЦІЯ

Гурмач В. В. Пошуку *in Silico* лігандів для доменів SH2. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата біологічних наук зі спеціальності 03.00.02 – біофізика. - Київський національний університет імені Тараса Шевченка МОН України, Київ, 2016.

Дисертація присвячена створенню цільової бібліотеки потенційних лігандів доменів SH2 та їх біологічній перевірці. Проаналізовано усі наявні домени SH2 та їх комплекси з нативними лігандами, зокрема виявлено 56 структур з унікальними амінокислотними послідовностями, які поділили на 6 груп та 8 репрезентативних кишени зв'язування. Встановлено чотири ключові точки зв'язування лігандів у кишени зв'язування доменів SH2. Отримано чотири моделі докінгу "домен SH2-ліганд": практично всюди ліганд щільно заповнює сайт зв'язування. Загалом отримали три бібліотеки потенційних лігандів доменів SH2, які містили 1015, 3018 і 10463 структур. Серед них виділили усі активні і неактивні речовини щодо доменів SH2 та побудували дев'ять унікальних фармакофорних моделей. Також побудували GTM модель передбачення активності потенційних лігандів. Виділили декілька різних хемотипів речовин, здатних імітувати фосфорильований тирозин. У біологічному тесті оцінили інгібуючий ефект потенційних лігандів доменів SH2 на проліферацію пухлинних клітин та кіназну активність Vtk та Syk кіназ людини.

Ключові слова: домен SH2, фармакофорна модель, GTM модель, Vtk та Syk кінази.

АННОТАЦИЯ

Гурмач В.В. Поиск *in Silico* лигандов для доменов SH2. - Рукопись.

Диссертация на соискание ученой степени кандидата биологических наук по специальности 03.00.02 - биофизика. - Киевский национальный университет имени Тараса Шевченко МОН Украины, Киев, 2016.

Диссертация посвящена созданию целевой библиотеки потенциальных лигандов доменов SH2 и их биологической проверке. Проанализированы все имеющиеся домены SH2 и их комплексы с нативными лигандами, в частности обнаружено 56 структур с уникальными аминокислотными последовательностями, которые разделили на 6 групп и 8 репрезентативных карманов связывания. Установлено четыре ключевые точки связывания лигандов в кармане связывания доменов SH2. Получено четыре модели докинга "домен SH2-лиганд": практически везде лиганд плотно заполняет сайт связывания. В целом получили три библиотеки потенциальных лигандов доменов SH2, содержащих 1015, 3018 и 10463 структур. Из активных вещества по отношению к доменам SH2 построили 9 уникальных фармакофорных моделей. Также построили GTM модель предсказания активности потенциальных лигандов. Выделили несколько различных хемотипов, способных имитировать фосфорилированный тирозин. При

биологической проверке оценили ингибирующий эффект потенциальных лигандов доменов SH2 на пролиферацию опухолевых клеток и киназную активность Btk и Syk киназ человека.

Ключевые слова: домен SH2, фармакофорная модель, GTM модель, Btk и Syk киназы.

ABSTRACT

Hurmach V.V. in Silico search of ligands for SH2 domain. - Manuscript.

Dissertation for the Candidate of Biological Sciences degree in speciality 03.00.02 - biophysics. - Taras Shevchenko National University of Kyiv, Ministry of Education and Science of Ukraine, Kyiv, 2016.

Dissertation is devoted to the library creation of potential ligands for SH2 domains and their biological verification. All available SH2 domains and their complexes with ligands were analyzed, 56 structures with unique amino acid sequences were identified and divided into 6 groups and 8 representative binding pockets. Comparable domain structures are characterized by high/average levels of conservation (group 1 - 92/91%, group 2 - 81/87%, group 3 - 56/62%, group 4 - 37/51%, group 5 - 50/68% and group 6 - 51/61%) and not significant Rmsd difference (0.8 Å, 1.16 Å, 1.6 Å, 1.9 Å, 2 Å and 1.4 Å). It is established that the binding site is more conservative (94/93%, 89/96%, 59/67%, 49/60%, 53/70% and 53/59%, respectively) and comprises 20-24 key amino acids: FLVRESETT (pTyr-binding part), β B-sheet; KHYKIR (central part), β D-sheet; ITSR and ADGLC (hydrophobic part of binding pocket), β G-fold and α B-helix.

Four key points of ligand binding pocket of SH2 domains were found: (1) H-bond acceptor (pTyr-binding part), (2) aromatic/hydrophobic part, (3) H-bond donor/acceptor, (4) hydrophobic part. Also, four "SH2-domain-ligand" docking models were received: almost in all cases ligand tightly fills in binding site. In general, three library of potential ligands to SH2 domains were obtained, which are containing 1015, 3018 and 10463 structures, respectively.

By using active compounds to SH2 domains 9 pharmacophore models were built. Their virtual screening showed good quality parameters, including the area under the ROC curve (AUC) varies between 75% and 100%. Thus, 1816 potentially active substances were selected. GTM model was created for predicting the activity of potential ligands too. Several different chemotypes, which can mimic phosphorylated tyrosine, were identified and as a result 1607 potentially active compounds were selected. In biological test the inhibitory effects of potential ligands of SH2 domains on proliferation of tumor cells and kinase activity of Btk and Syk kinases were evaluated.

Key words: SH2 domain, pharmacophore model, GTM model, Btk and Syk kinases.